Catherine Yeh

catherineyeh@g.harvard.edu | linkedin.com/in/catherinesyeh | catherinesyeh.github.io

EDUCATION

Harvard University *Cambridge, MA*
Ph.D. in Computer Science **2022 – Present**
Research Interests: Visualization, Interpretability, Human-AI Interaction
Advised by Martin Wattenberg & Fernanda Viégas

Williams College *Williamstown, MA*
B.A. in Computer Science & Cognitive Science **2018 – 2022**
Thesis: "Toward an Empirical Framework for Post-hoc Explainable AI"
Advised by Iris Howley
Graduated Summa Cum Laude (top 2% of class) with Highest Honors in Computer Science

INDUSTRY RESEARCH EXPERIENCE

Apple Research – Human-Centered Machine Intelligence Group *Seattle, WA*
Research Intern (*Mentors:* Fred Hohman, Donghao Ren, Yannick Assogba) **Summer 2024**

- Developed a human-in-the-loop data augmentation tool for increasing data diversity in unstructured text datasets using sparse autoencoders, embedding interpolation & visualization, and large language models.
- Evaluated interactive prototype on a real-world model safety task with professional red teamers [P7].

Microsoft Research – Human Understanding & Empathy Group *Redmond, WA*
Research Intern (*Mentors:* Gonzalo Ramos, Rachel Ng) **Summer 2023**

- Designed and implemented a personalized environment for collaborative writing with large language models (LLMs).
- Evaluated interactive prototype with user study to better understand how LLMs can assist and augment people's writing capabilities while preserving agency and ownership [P6].

Adobe Research – Media Intelligence Lab *San Jose, CA*
Research Scientist Intern (*Mentors:* Franck Dernoncourt, Nedim Lipka) **Summer 2022**

- Developed UI prototypes showcasing vision for the next-gen document reader [P2].
- Designed and implemented a novel NLP-powered acronym glossary for Adobe.

Microsoft Research – Productivity & Intelligence Group *Redmond, WA*
Undergraduate Research Intern (*Mentors:* Jenna Butler, Christian Bird) **Summer 2021**

- Built Microsoft Teams bot, personalized dashboards, and ML auto-coding system for qualitative survey responses.
- Analyzed personas and designed scalable, automated self-reflection interventions for study on hybrid workforce productivity and well-being [P1].

ACADEMIC RESEARCH EXPERIENCE

Harvard University – Insight & Interaction Lab *Allston, MA*
Computer Science Research Assistant (*Advisors:* Martin Wattenberg, Fernanda Viégas) **2022 - Present**

- Developing visualization-driven interfaces and techniques to support model interpretability and enable more productive, safe, and creative human-AI interactions.
- **Story Ribbons (in progress):** reimagining story visualization techniques with large language models.
- **TalkTuner:** developed a dashboard to provide model transparency and control of conversational language models through surfacing their internal "user models" [P5].
- **AttentionViz:** designed a new technique to visualize self-attention patterns in language & vision transformer models using joint query-key embeddings [P4].

Williams College – Computation² Lab *Williamstown, MA*
Computer Science Research Assistant (*Advisor:* Molly Feldman) **2021 - 2022**

- Studied state of replicability in human-computer interaction and developed set of guidelines to aid researchers in designing and publishing studies for replicability.

Williams College – Human-AI Interaction Lab *Williamstown, MA*
Computer Science Research Assistant (*Advisor:* Iris Howley) **2019 – 2022**

- Built interactive tutors for Bayesian Knowledge Tracing, an AI algorithm that predicts skill mastery.
- Developed a novel, evidence-based framework for explainable AI using cognitive task analysis, user-centered design, and learning theory [P3].

Williams College – Concepts & Categories Lab *Williamstown, MA*
Cognitive Science Research Assistant (*Advisor:* Safa Zaki) **2018 – 2019**

- Designed, executed, and analyzed eye-tracking experiments to study the relationship between active / passive learning and interleaving / blocking effects in categorization tasks.

PUBLICATIONS

[P7] **Yeh, C.,** Ren, D., Assogba, Y., Moritz, D., & Hohman, F. (In submission). Exploring Empty Spaces: Human-in-the-Loop Data Augmentation. *Preprint:* arxiv.org/abs/2410.01088

[P6] **Yeh, C.,** Ramos, G., Ng, R., Huntington, A., & Banks, R. (In submission). GhostWriter: Augmenting Collaborative Human-AI Writing Experiences Through Personalization and Agency. *Preprint:* arxiv.org/abs/2402.08855

[P5] Chen, Y., Wu, A., DePodesta, T., **Yeh, C.,** Li, K., Castillo Marin, N., Patel, O., Riecke, J., Raval, S., Seow, O., Wattenberg, M., & Viégas, F. (In submission). Designing a Dashboard for Transparency and Control of Conversational AI. *Preprint:* arxiv.org/abs/2406.07882

[P4] **Yeh, C.,** Chen, Y., Wu, A., Chen, C., Viégas, F., & Wattenberg, M. (2023). AttentionViz: A Global View of Transformer Attention. *IEEE Visualization Conference.* ieeexplore.ieee.org/document/10297591

[P3] **Yeh, C.,** Cowit, N., & Howley, I. (2023). Designing for Student Understanding of Learning Analytics Algorithms. *International Conference on Artificial Intelligence in Education.* link.springer.com/chapter/10.1007/978-3-031-36272-9_43

[P2] **Yeh, C.,** Dernoncourt, F., & Lipka, N. (2023). Envisioning the Next-Gen Document Reader. *AAAI Workshop on Scientific Document Understanding.* doi.org/10.48550/arXiv.2302.07492

[P1] Butler, J., & **Yeh, C.** (2022). Walk a Mile in Their Shoes: The Covid Pandemic Through the Lens of Four Tech Workers. *Communications of the ACM.* dl.acm.org/doi/10.1145/3561989
Also featured in *ACM Queue:* dl.acm.org/doi/10.1145/3534860

TEACHING EXPERIENCE

Harvard University **2023 - Present**
Research Topics in Human-Computer Interaction (COMPSCI 279R)

Williams College **2019 – 2022**
Principles of Programming Languages (CSCI 334), Computational Biology (CSCI 315), Algorithm Design & Analysis (CSCI 256), Data Structures & Advanced Programming (CSCI 136)

OTHER INDUSTRY EXPERIENCE

Oracle Cloud Infrastructure (OCI)
Software Engineer Intern, *Summer 2020*

Sunshine
Software Engineer Intern, *Winter 2020*

HONORS & AWARDS

2024 National Science Foundation Graduate Research Fellow
Kempner Institute Graduate Fellow

2022 Sigma Xi International Scientific Research Honor Society Inductee
CRA Outstanding Undergraduate Award Runner Up

2021 NCWIT Collegiate Award Finalist
Adobe Research Women-in-Technology Scholarship Finalist
Phi Beta Kappa National Honor Society Junior Year Inductee (top 5% of class)
Grace Hopper Conference ACM Student Research Competition Finalist
Williams College Ward Prize Finalist

2019 Williams College Computer Science Class of 1960s Scholar (2019 – 2022)
National USCLAP Statistics Competition 3rd Place Winner
Williams College Summer Science Research Fellow
Grace Hopper Conference Scholar

2018 Williams College Dean's List (2018 – 2022)

SERVICE

Mentor – Research
Vicki Xu, *Harvard '23:* logit lens on vision transformer
Cynthia Chen, *Harvard '24:* transformer interpretability

Mentor – Other
Harvard SEAS Research Program (2024 – Present)
Harvard Women in STEM (2022 – Present)
Williams Underrepresented Identities in CS (2020-2022)

Program Committee
IEEE VISxAI Workshop (2024)

Reviewer
ACM CHI (2024), ACM DIS (2024)

Volunteer
IEEE VIS Student Volunteer (2023)